

Appln. No. 09/994,396
Amendment dated Mar. 10, 2005
Reply to Office Action of Oct. 22, 2004
Docket No. DE9-2000-0096 (270)

This listing of claims will replace all prior versions and listings of claims in the instant application:

LISTING OF CLAIMS

1. (Currently Amended) A method of ~~improving speech recognition~~ automatically training a word database used to by a speech recognition engine to speech-to-text convert utterances comprising:

taking a realization of spoken audio and a first representation that is an allegedly true textual representation for said realization;

performing a speech recognition on said realization using the word database thereby producing a second representation that is a textual representation for said realization;

aligning said first representation and said second representation on a word-by-word basis;

detecting at least one word aligned in the aligning step that is different in the first representation and the second representation; and

~~selecting single words from said first representation and corresponding aligned single words from said second representation and pairing said aligned single words, wherein said first and said second representations are different; and~~

automatically updating [[a]] the word database to include the at least one detected word using said selected paired words together with said corresponding speech segment extracted from the aligned realization, wherein the updating step extends the word database.

{WP215417;3}

Appln. No. 09/994,396
Amendment dated Mar. 10, 2005
Reply to Office Action of Oct. 22, 2004
Docket No. DE9-2000-0096 (270)

2. (Currently Amended) The method according to claim 1, wherein said ~~selecting~~ detecting step uses speech recognition information derived from said speech recognition.
3. (Original) The method according to claim 2, wherein said aligning step reveals time information corresponding to the alignment between said realization and said first representation.
4. (Currently Amended) The method according to claim ~~[[2]]~~ 1, said updating step further comprising:
comparing the recognition quality of said speech recognition of said realization with the recognition quality of a corresponding single word entry existing in said word database.
5. (Original) The method according to claim 4, wherein said first and said second representations are comprised of segments, said comparing step further comprising: tagging said segments of said first and said second representations where both said first and said second representations consist of a single word.
6. (Currently Amended) A method of improving speech recognition comprising:
taking a realization of spoken audio and a first representation that is an allegedly true textual representation for said realization;
performing a speech recognition on said realization using a word database thereby producing a second representation that is a textual representation for said realization;

{WP215417;3}

Appln. No. 09/994,396
Amendment dated Mar. 10, 2005
Reply to Office Action of Oct. 22, 2004
Docket No. DE-2000-0096 (270)

aligning said first representation and said second representation;
detecting matching word pairs resulting from said aligning step; and
~~selecting single words from said first representation and corresponding aligned~~
~~single words from said second representation and pairing said aligned single words,~~
~~wherein said first and said second representations are identical; and~~
automatically updating a pronunciation database the word database using said
~~selected paired words~~ matching word pairs together with [[said]] corresponding
[[aligned]] realization information, wherein the updating step extends the word database.

7. (Currently Amended) The method according to claim 6, wherein said ~~selecting~~ detecting step uses speech recognition information derived from said speech recognition.
8. (Original) The method according to claim 7, wherein said aligning step reveals time information corresponding to the alignment between said realization and said first representation.
9. (Currently Amended) The method according to claim ~~[[7]]~~ 6, said updating step further comprising:
comparing the recognition quality of said speech recognition of said realization with the recognition quality of a corresponding single word entry existing in said pronunciation database.

{WP215417;3}

Appln. No. 09/994,396
Amendment dated Mar. 10, 2005
Reply to Office Action of Oct. 22, 2004
Docket No. DE9-2000-0096 (270)

10. (Original) The method according to claim 9, wherein said first and said second representations are comprised of segments, said comparing step further comprising:

{WP215417;3}

tagging said segments of said first and said second representations where both said first and said second representations consist of a single word.

11. (Currently Amended) A system for improving speech recognition of a speech recognizer, said system comprising:

a word database used by the speech recognizer to perform speech recognition tasks;

an aligner configured to align a first representation and a second representation produced by said speech recognizer;

a classifier configured to compare said aligned first representation with said aligned second representation; and

a selector configured to select corresponding single word pairs from said aligned first representation and said aligned second representation, wherein at least one word of the single word pairs selected by the selector are combined with a realization of spoken audio from which said speech recognizer produced the second representation, the combination being used to update the word database, wherein the updating step extends the word database.

12. (Currently Amended) The system according to claim 11, wherein said first representation and said second representation are different, and wherein updating the word database results in new words being added to the word database.

13. (Original) The system according to claim 11, wherein said first representation and said second representation are identical.

(WP215417.3)

Certificate Under 37 CFR 1.8(a)

I hereby certify that this correspondence is being transmitted via facsimile to MAILSTOP AF,
Commissioner for Patents, Arlington, VA at 703-872-9306 on

Date

_____, Reg. No. 52,667
Brian K. Buchheit, Esquire

U.S. Appl. No. 09/994,396
Amendment Date: March 08, 2005
Reply to Office Action of October 22, 2004
Docket No. DE9-2000-0096 (270)

14. (Currently Amended) The system according to claim 11, further comprising means for updating ~~a word database or~~ a pronunciation database based upon the combination using single word pairs selected by said selector.

15. (Original) The system according to claim 11, said aligner further comprising:
means for generating time information corresponding to time alignment between said first representation and said second representation.

16. (Original) The system according to claim 15, wherein said first and said second representations comprise segments, said classifier further comprising:
means for tagging said segments of said first representation and said second representation where said first representation and said second representation consist of a single word.

17. (Currently Amended) A machine-readable storage, having stored thereon a computer program having a plurality of code sections executable by a machine for causing the machine to perform the steps of:

taking a realization of spoken audio and a first representation that is an allegedly true textual representation for said realization;

performing a speech recognition on said realization using the word database thereby producing a second representation that is a textual representation for said realization;

aligning said first representation and said second representation on a word-by-word basis;

detecting at least one word aligned in the aligning step that is different in the first representation and the second representation: and

{WP215417;3}

U.S. Appl. No. 09/994,396
Amendment Dated March 08, 2005
Reply to Office Action of October 22, 2004
Docket No. DE9-2000-0096 (270)

~~selecting single words from said first representation and corresponding aligned single words from said second representation and pairing said aligned single words, wherein said first and said second representations are different; and~~

automatically updating [[a]] the word database to include the at least one detected word using said selected paired words together with said corresponding speech segment extracted from the aligned realization, wherein the updating step extends the word database.

18. (Currently Amended) The machine-readable storage according to claim 17, wherein said ~~selecting~~ detecting step uses speech recognition information derived from said speech recognition.

19. (Original) The machine-readable storage according to claim 18, wherein said aligning step reveals time information corresponding to the alignment between said realization and said first representation.

20. (Currently Amended) The machine-readable storage according to claim ~~[[18]]~~ 17, said updating step further comprising:

comparing the recognition quality of said speech recognition of said realization with the recognition quality of a corresponding single word entry existing in said word database.

21. (Original) The machine-readable storage according to claim 20, wherein said first and said second representations are comprised of segments, said comparing step further comprising:

tagging said segments of said first and said second representations where both said first and said second representations consist of a single word.

{WP215417:3}

U.S. Appln. No. 09/994,396
Amendment Date: March 08, 2005
Reply to Office Action of October 22, 2004
Docket No. DE9-2000-0096 (270)

22. (Currently Amended) A machine-readable storage, having stored thereon a computer program having a plurality of code sections executable by a machine for causing the machine to perform the steps of:

taking a realization of spoken audio and a first representation that is an allegedly true textual representation for said realization;

performing a speech recognition on said realization using a word database thereby producing a second representation that is a textual representation for said realization;

aligning said first representation and said second representation;

detecting matching word pairs resulting from said aligning step; and

~~selecting single words from said first representation and corresponding aligned single words from said second representation and pairing said aligned single words, wherein said first and said second representations are identical, and~~

automatically updating a pronunciation database the word database using said selected paired words matching word pairs together with [[said]] corresponding aligned realization information, wherein the updating step extends the word database.

23. (Currently Amended) The machine-readable storage according to claim 22, wherein said ~~selecting~~ detecting step uses speech recognition information derived from said speech recognition.

24. (Original) The machine-readable storage according to claim 23, wherein said aligning step reveals time information corresponding to the alignment between said realization and said first representation.

25. (Currently Amended) The machine-readable storage according to claim ~~[[23]]~~22, said updating step further comprising:

(WP215417;3)

U.S. Appln. No. C9/994,396
Amendment Date: March 08, 2005
Reply to Office Action of October 22, 2004
Docket No. DE9-1000-0096 (270)

comparing the recognition quality of said speech recognition of said realization with the recognition quality of a corresponding single word entry existing in said pronunciation database.

26. (Original) The machine-readable storage according to claim 25, wherein said first and said second representations are comprised of segments, said comparing step further comprising:

tagging said segments of said first and said second representations where both said first and said second representations consist of a single word.

27. (New) The method of claim 1, further comprising the step of:
obtaining said first representation using an optical character recognition technology.

28 (New) The method of claim 1, wherein the word database is a speaker dependent database used to adapt the speech recognition to a particular speaker.

{WP215417;3}